2007

Adaptive reorganization of database structures through dynamic vertical partitioning of relational tables

Liu Zhenjie

University of Wollongong

Adaptive Reorganization of Database Structures Through Dynamic Vertical Partitioning of Relational Tables

A thesis submitted in fulfillment of the
requirements for the award of the degree

Master of Computer Science by Research

from

UNIVERSITY OF WOLLONGONG

by

Zhenjie Liu

School of Information Technology and Computer Science
October 2007

Dedicated to
My parents and wife Yanting

Declaration

This is to certify that the work reported in this thesis was done by the author, unless specified otherwise, and that no part of it has been submitted in a thesis to any other university or similar institution.

Zhenjie Liu
October 23, 2007

Publications

Zhenjie Liu and Janusz R. Getta. Optimization of query processing through constrained vertical partitioning of relational tables. In *DBA'06: Proceedings of the 24th IASTED international conference on Database and applications*, pages 221-227, Anaheim, CA, USA, 2006. ACTA Press.

Abstract

Performance tuning of relational database systems is always a challenging task for database administrator. Automated performance tuning has been proposed recently as a new approach to detect and to eliminate performance problems and to support the decisions of database administrators.

This work considers one of the techniques used in automated performance tuning, dynamic vertical partitioning. Dynamic vertical partitioning of relational tables is one of the ways in which the physical structures of a relational database can be reorganised automatically in order to improve the performance of future database applications. The thesis presents how dynamic vertical partitioning can be used for the comprehensive analysis and optimisation of an adaptive reorganisation of database structures. In particular, we propose the algorithms to use to predict the future workload of the system, to analyse the characteristics of the workload, and to find a near optimal vertical partitioning of relational tables. Then, we discuss the implementation aspects of vertical partitioning with the materialized view and index-based techniques.

Our contributions to automated performance tuning of relational database systems can be summarised as follows:

1. Propose a cost model to perform a detailed analysis of the costs of query and data manipulation processing over a given configuration of a relational database;

2. Propose a new algorithm for vertical partitioning of relational schemas in a database system with a given level of redundancies for a given workload;

3. Discuss the limitations of static vertical partitioning and propose dynamical vertical partitioning;

4. Discuss the characteristics of workload in order to predict the future workloads of the system;

5. Implementation aspects of vertical partition discussion: materialized view based and index based;

6. Discussion of the implementation of a vertical partition as a virtual view;

7. Conduct experiments to confirm the correctness of the cost model used by the vertical partitioning algorithm and demonstrate the expected performance gains from the partitioning

Acknowledgements

The research work for this thesis was undertaken at the University of Wollongong.

Firstly, I would like to thank my supervisor, Dr. Janusz R.Getta for his guidance and help in my research. Without him, this thesis would not have been possible.

Also, I wish to acknowledge the support I have received from all the staff in the School of IT & CS, University of Wollongong.

Finally, I would like to thank my family and friends for their enduring love and support.

Contents

List of Tables

List of Figures